WHAT IS CLAIMED IS:

1. A method for filtering an input text stream, the method comprising:

receiving a definition of a filter configuration; and

modifying the input text stream according to the filter configuration so as to generate a filtered text stream, the filtered text stream including positioning information for the input text stream.

2. The method as recited in claim 1 wherein the receiving a definition of the filter configuration

includes:

receiving a definition of a plurality of patterns; and

receiving a definition of a respective association between each of the plurality of patterns and a respective executable action.

3. The method as recited in claim 2 wherein the receiving the definition of the filter

configuration further includes processing each of the plurality of patterns and the respective association so as to form a scanner data structure capable of comparing the input text stream to at least one of the plurality of patterns and causing execution of the associated executable action upon a match of the input text stream with the respective one of the plurality of patterns.

4. The method as recited in claim 3 wherein the processing and the comparing are performed in a same active process.

5. The method as recited in claim 2 wherein at least one of the respective executable action includes replacing a matched text in the input text stream with a respective replacement text.

6. The method as recited in claim 5 wherein the replacement text is determined dynamically.

7. The method as recited in claim 6 wherein the replacement text is computed by the respective executable action.

8. The method as recited in claim 5 wherein the replacement text is determined using at least a portion of the matched text.

9. The method as recited in claim 1 wherein modifying the input text stream includes replacing at least one character in the input text stream with a replacement text.

10. The method as recited in claim 8 wherein the replacement text is determined dynamically.

11. The method as recited in claim 9 wherein the replacement text is determined using at least a portion of the matched text.

12. The method as recited in claim 8 further comprising comparing the input text stream to a plurality of patterns and causing the replacing upon a match of the input text stream with the respective one of the plurality of patterns.

13. The method as recited in claim 8 wherein the modifying the input text stream includes determining a difference in a number of characters between the at least one character and the replacement text.

14. The method as recited in claim 13 wherein the positioning information includes the difference in the number of characters between the at least one character and the replacement text.

15. The method as recited in claim 1 wherein the positioning information refers to positions of characters in the input text stream.

16. The method as recited in claim 1 wherein the positioning information includes directives.

17. The method as recited in claim 16 wherein the directives include executable actions, the executable actions being executable by a scanner receiving the filtered text stream.

18. The method as recited in claim 1 further comprising providing a modified scanner generator skeleton configured to receive the filtered text stream, the modified scanner generator skeleton being useable to generate a scanner capable of applying the positioning information.

19. The method as recited in claim 18 further comprising generating the scanner using the modified scanner generator skeleton.

20. The method as recited in claim 19 wherein the scanner forms part of a parser.

21. The method as recited in claim 20 wherein the part includes a front end of the parser.

22. The method as recited in claim 3 wherein the receiving a definition of the filter configuration further includes:

associating at least one respective start state of a plurality of start states with each of the plurality of patterns; and

setting a current start state, the current start state being one of the plurality of start states; and

wherein the processing includes processing the at least one respective start state

along with each associated pattern so as to form a part of the scanner data structure, the comparing being performed so as to compare only the patterns of the plurality of patterns associated with a respective start state equal to the current start state.

23. The method as recited in claim 22 wherein each respective start state is included in a context.

24. An apparatus for filtering a text stream, the apparatus comprising:

a filter configuration data structure; and

a scanner unit configured to modify the input text stream using the filter configuration data structure so as to generate a filtered text stream, the filtered text stream including positioning information for the input text stream.

25. The apparatus as recited in claim 24 wherein the filter configuration data structure includes:

a plurality of patterns; and

a respective association between each of the plurality of patterns and a respective executable action.

26. The apparatus as recited in claim 25 wherein the filter configuration is formed by processing each of the plurality of patterns and the respective association so as to form a scanner data structure capable of comparing the input text stream to at least one of the plurality of patterns and causing execution of the associated executable action upon a match of the input text stream with the respective one of the plurality of patterns.

27. The apparatus as recited in claim 26 wherein the processing and the comparing are performed in a same active process.

28. The apparatus as recited in claim 25 wherein at least one of the respective

executable action includes replacing a matched text in the input text stream with a respective replacement text.

29. The apparatus as recited in claim 28 wherein the replacement text is determined dynamically.

30. The apparatus as recited in claim 29 wherein the replacement text is computed by the respective executable action.

31. The apparatus as recited in claim 28 wherein the replacement text is determined using at least a portion of the matched text.

32. The apparatus as recited in claim 24 wherein modifying the input text stream includes replacing at least one character in the input text stream with a replacement text.

33. The apparatus as recited in claim 32 wherein the replacement text is determined dynamically.

34. The apparatus as recited in claim 32 wherein the replacement text is determined using at least a portion of the matched text.

35. The apparatus as recited in claim 32 wherein the scanner unit is further configured to

compare the input text stream to a plurality of patterns and causing the replacing upon a match of the input text stream with the respective one of the plurality of patterns.

36. The apparatus as recited in claim 32 wherein modifying the input text stream further includes determining a difference in a number of characters between the at least one character and the replacement text.

37. The apparatus as recited in claim 36 wherein the positioning information includes the difference in the number of characters between the at least one character and the replacement text.

38. The apparatus as recited in claim 24 wherein the positioning information refers to positions of characters in the input text stream.

39. The apparatus as recited in claim 24 wherein the positioning information includes directives.

40. The apparatus as recited in claim 39 wherein the directives include executable actions, the executable actions being executable by a scanner receiving the filtered text stream.

41. The apparatus as recited in claim 24 further comprising a modified scanner generator skeleton configured to receive the filtered text stream, the modified scanner generator skeleton being useable to generate a scanner capable of applying the positioning information.

42. The apparatus as recited in claim 41 further comprising a scanner, the scanner being generated using the modified scanner generator skeleton.

43. The apparatus as recited in claim 42 wherein the scanner forms a part of a parser.

44. The apparatus as recited in claim 43 wherein the part includes a front end of the parser.

45. The apparatus as recited in claim 26 wherein the filter configuration further includes:

at least one respective start state of a plurality of start states associated with each of the plurality of patterns; and

setting a current start state, the current start state being one of the plurality of start states; and

wherein the processing includes processing the at least one respective start state along with each associated pattern so as to form a part of the scanner data structure, the comparing being performed so as to compare only the patterns of the plurality of patterns associated with a respective start state equal to the current start state.

46. The apparatus as recited in claim 24 wherein each respective start state is included in a context.

47. A computer readable medium having stored thereon computer executable process steps operative to perform a method for filtering an input text stream, the method comprising:

receiving a definition of a filter configuration; and

modifying the input text stream according to the filter configuration so as to generate a filtered text stream, the filtered text stream including positioning information for the input text stream.

48. The computer readable medium as recited in claim 47 wherein the definition of the filter configuration is received from a data structure.

49. A parsing device comprising:

a filter configuration data structure;

a first scanner unit configured to modify the input text stream using the filter configuration data structure so as to generate a filtered text stream, the filtered text

stream including positioning information for the input text stream; and

a second scanning unit configured to process the filtered text stream so as to apply the positioning information.

50. The parsing device as recited in claim 49 wherein the modifying the input text stream includes replacing at least one character in the input text stream with a replacement text and determining a difference in the number of characters between the at least one character and the replacement string, and wherein the positioning information includes the difference.

51. The parsing device as recited in claim 49 wherein the positioning information includes directives.

52. The parsing device as recited in claim 51 wherein the directives include executable actions, the executable actions being executable by the second scanning unit.

53. The parsing device as recited in claim 49 wherein the second scanning unit includes a modified scanner generator skeleton configured to receive the filtered text stream, the modified scanner generator skeleton being useable to generate a scanner capable of applying the positioning information.